



File Copy  
09/982223

# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 161303

TO: David Lamberston  
Location: REM/2B79/2C70  
Art Unit: 1636  
Tuesday, August 09, 2005

Case Serial Number: 09/982223

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Lamberston,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



STIC-Biotech/ChemLib

161303

**From:** Lambertson, David  
**Sent:** Tuesday, August 02, 2005 3:18 PM  
**To:** Lambertson, David; STIC-Biotech/ChemLib  
**Subject:** RE: Search Request

Sorry, I mistakenly removed the Default search request. Please search SEQ ID NO: 1 and 2 against the nucleotide databases for:

1. Default Search.
2. Interference Search.

Thanks.

-----Original Message-----

**From:** Lambertson, David  
**Sent:** Tuesday, August 02, 2005 12:33 PM  
**To:** STIC-Biotech/ChemLib  
**Cc:** Lambertson, David  
**Subject:** Search Request

## Search Request

<b>Examiner's Name:</b>	David Lambertson
<b>Examiner #:</b>	79514
<b>Art Unit:</b>	1636
<b>Room #:</b>	02B79-Remsen
<b>Mailbox room#:</b>	02C70-Remsen
<b>Phone:</b>	(571) 272-0771
<b>Results Format:</b>	paper

**Serial #:09/982,223**

**Please Search:**

**Nucleic Acid databases for:**

**SEQ ID No: 1 and 2**

**Including:**

1. Interference Search

8/2/05  
2-NA  
Delt  
8/9/05

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 08:47:41 ; Search time 8137.61 Seconds  
(without alignments)  
19369.933 Million cell updates/sec

Title: US-09-982-223A-2  
Perfect score: 4141  
Sequence: 1 aatgaagaccacactgtca.....agtcctcagaaaaagg999g 4141

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gest1:\*  
9: gb\_gest2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	672.4	16.2	865	7	CK125894	CK125894 BES182411
C 2	666	16.1	844	8	BZ574513	BZ574513 msh2_3706
C 3	662	16.0	811	9	ATH517156	AJ517156 Arabidops
C 4	661.4	16.0	800	1	AJ281449	AJ281449 4A3A-P4D5
C 5	659.4	15.9	734	7	CO478328	CO478328 GQ0131.B7
C 6	656	15.8	700	5	BQ660293	BQ660293 H01N05W
C 7	654	15.8	1004	1	AJ281480	AJ281480 4A3A-P4G8
C 8	646.8	15.6	1048	7	CO523236	CO523236 Acty4_50
C 9	642.4	15.5	846	7	CV468077	CV468077 est_1_van
C 10	642	15.5	691	6	CB864071	CB864071 HH07007Y
C 11	638	15.4	700	6	CA305674	CA305674 EST000022
C 12	636	15.4	859	8	BZ574002	BZ574002 msh2_3467
C 13	630.8	15.2	1208	8	BZ577775	BZ577775 msh2_5567
C 14	626.8	15.1	966	8	BZ570738	BZ570738 msh2_1513
C 15	626.4	15.1	966	8	BZ570738	BZ570738 msh2_1513
C 16	626	15.1	1070	5	BP875099	BP875099 BP875099
C 17	624	15.1	1070	5	BP874724	BP874724 BP874724
C 18	623.8	15.1	854	8	BZ570648	BZ570648 msh2_1469
C 19	623.4	15.1	733	8	BZ048997	BZ048997 Jur52g07.
C 20	623.4	15.1	1340	8	BZ572620	BZ572620 msh2_2721
C 21	623.4	15.1	1574	8	BZ572566	BZ572566 msh2_2693
C 22	621.4	15.0	1404	8	BZ572478	BZ572478 msh2_2653
C 23	621	15.0	665	6	CB873278	CB873278 HCL1H05Y
C 24	620	15.0	754	8	BZ040933	BZ040933 lka28c05.

```
25 618.6 14.9 695 8 BH929757
26 618.4 14.9 658 8 BH970618
27 618.4 14.9 667 8 BH964466
28 618.4 14.9 682 8 BH930868
29 618.4 14.9 684 8 BH955957
30 618.4 14.9 687 8 BH969248
31 618.4 14.9 690 8 BH921350
32 618.4 14.9 690 8 BZ088472
33 618.4 14.9 692 8 BH964078
34 618.4 14.9 692 8 BH972004
35 618.4 14.9 692 8 BH978479
36 618.4 14.9 693 8 BH960007
37 618.4 14.9 695 8 BH947144
38 618.4 14.9 697 8 BH934795
39 618.4 14.9 697 8 BH949519
40 618.4 14.9 697 8 BZ088379
41 618.4 14.9 703 8 BH934948
42 618.4 14.9 703 8 BH938426
43 618.4 14.9 703 8 BH953754
44 618.4 14.9 709 8 BH974652
45 618.4 14.9 710 8 BH984254
```

## ALIGNMENTS

```
RESULT 1
CK125894/c
LOCUS CK125894 865 bp mRNA linear EST 01-MAR-2004
DEFINITION BSM182411A2B03 BES1824 Hordeum vulgare subsp. vulgare cDNA clone
MPMGp2010B0311 5-PRIME, mRNA sequence.
ACCESSION CK125894
VERSION CK125894.1 GI:44808896
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 865)
Kramer, A., Feilner, T., Possling, A., Radchuk, V., Weschke, W.,
Buerkle, L. and Kersten, B.
Application of the protein microarray technology for the
identification of expression library derived target proteins for
barley protein kinase CK2
Unpublished (2003)
Contact: Birgit Kersten* and Winfriede Weschke**
*Plant Protein Chip Group, Department Lehrach, **Department
Molecular Genetics, Gene Expression Group
**Max-Planck-Institute for Molecular Genetics, **Institute of Plant
Genetics and Crop Plant Research Gatersleben
*Thuenstr. 73, D-14195 Berlin, Germany, **Corrensstrasse 3, D-06466
Gatersleben, Germany
Tel: **49 (0) 30/84131648, **49 (0) 394825500
Fax: **49 (0) 30/84131128, **49 (0) 394825237
Email: *kersten@molgen.mpg.de, **weschke@ipk-gatersleben.de
Insert Length: 865 Std Error: 0.00
Place: 11 row: B column: 3
Seq primer: PQE65.
Location/Qualifiers
1. .865
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:946193"
/db_xref="taxon:112509"
/clone="MPMGp2010B0311"
/tissue_type="embryoosac"
/dev_stages="0-10 DAF (days after flowering)"
/lab_host="E. coli, SCS-1/pSE111"
/clone_lib="BES1824"
/note="Vector: PQE30NST (AF074376); Site_1: Salt; Site_2:
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 01:28:50 ; Search time 1577.44 Seconds  
(without alignments)  
17016.962 Million cell updates/sec

Title: US-09-982-223A-2

Perfect score: 4141

Sequence: 1 aatgaagacccacactgta.....agctccagaaaaagggggg 4141

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_PUB\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10J\_PUBCOMB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4141	100.0	4141	10	US-09-982-223A-2
2	2131.6	51.5	5713	9	US-09-963-206B-5
3	2131.6	51.5	5713	9	US-09-966-976A-5
4	2131.6	51.5	5713	11	US-09-963-247A-5
5	1840	44.4	5782	10	US-09-982-223A-1
6	1825	44.1	6825	22	US-10-987-388-36
7	1825	44.1	6851	22	US-10-987-388-35

44.1	1825	44.1	10100	9	US-09-208-827-1	Sequence 1, Appli
44.1	1825	44.1	10100	14	US-10-043-074-1	Sequence 1, Appli
6870	21	43.5	6870	21	US-10-833-951-2	Sequence 2, Appli
7654	13	40.9	7654	13	US-10-006-773-1	Sequence 1, Appli
7654	13	40.9	7654	13	US-10-006-771A-1	Sequence 1, Appli
8614	9	38.6	8614	9	US-09-208-827-5	Sequence 5, Appli
8614	14	38.6	8614	14	US-10-043-074-5	Sequence 5, Appli
963	21	38.5	963	21	US-10-913-288-1	Sequence 1, Appli
6219	9	38.3	6219	9	US-09-963-206B-4	Sequence 4, Appli
6219	9	38.3	6219	9	US-09-966-976A-4	Sequence 4, Appli
6219	11	38.3	6219	11	US-09-963-247A-4	Sequence 4, Appli
6620	8	36.2	6620	8	US-08-786-531B-3	Sequence 3, Appli
5363	8	36.1	5363	8	US-08-786-531B-2	Sequence 2, Appli
8282	9	35.2	8282	9	US-09-963-206B-7	Sequence 7, Appli
8282	9	35.2	8282	9	US-09-966-976A-7	Sequence 7, Appli
8282	11	35.2	8282	11	US-09-963-247A-7	Sequence 7, Appli
8345	9	35.2	8345	9	US-09-963-206B-8	Sequence 8, Appli
8345	11	35.2	8345	11	US-09-966-976A-8	Sequence 8, Appli
8345	11	35.2	8345	11	US-09-963-247A-8	Sequence 8, Appli
6254	16	35.2	6254	16	US-10-317-078-1	Sequence 1, Appli
6254	21	35.2	6254	21	US-10-789-938B-1	Sequence 1, Appli
4922	9	34.9	4922	9	US-09-963-206B-6	Sequence 6, Appli
4922	9	34.9	4922	9	US-09-966-976A-6	Sequence 6, Appli
4922	11	34.9	4922	11	US-09-963-247A-6	Sequence 6, Appli
8388	9	34.7	8388	9	US-09-987-601-1	Sequence 1, Appli
8889	18	34.5	8889	18	US-10-677-558-4	Sequence 4, Appli
11364	18	34.5	11364	18	US-10-677-558-11	Sequence 11, Appli
11394	18	34.5	11394	18	US-10-677-558-5	Sequence 5, Appli
5174	21	34.3	5174	21	US-10-954-645-7	Sequence 7, Appli
6250	17	29.9	6250	17	US-10-182-329-109	Sequence 109, Appli
6250	18	29.9	6250	18	US-10-182-327-16	Sequence 16, Appli
8161	9	29.7	8161	9	US-09-759-152-8	Sequence 8, Appli
6700	9	29.7	6700	9	US-09-759-152-1	Sequence 1, Appli
8175	9	29.7	8175	9	US-09-759-152-4	Sequence 4, Appli
8518	9	29.7	8518	9	US-09-759-152-4	Sequence 4, Appli
7840	10	29.6	7840	10	US-09-957-458B-5	Sequence 5, Appli
8852	10	29.6	8852	10	US-09-957-458B-6	Sequence 6, Appli
9772	22	23.5	9772	22	US-10-763-479-2	Sequence 2, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-982-223A-2  
; Sequence 2, Application US/09982223A  
; Publication No. US20030175972A1  
; GENERAL INFORMATION:  
; APPLICANT: Daley, George O.  
; APPLICANT: Koh, Eugene Y.  
; TITLE OF INVENTION: EXPRESSION VECTORS AND USES THEREOF  
; FILE REFERENCE: 13086-002001  
; CURRENT APPLICATION NUMBER: US/09/982,223A  
; CURRENT FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/241,879  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4141  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated nucleic acid  
US-09-982-223A-2

Query Match 100.0%; Score 4141; DB 10; Length 4141;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AATGAAGACCCACCTGTAGTTTGGCAACTAGCGCGCGCATTAACCTTCGTATAGCA 60  
DB 1 AATGAAGACCCACCTGTAGTTTGGCAACTAGCGCGCGCATTAACCTTCGTATAGCA 60

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 10:56:46 ; Search time 425.242 Seconds  
(without alignments)  
15934.027 Million cell updates/sec

Title: US-09-982-223A-2  
Perfect score: 4141  
Sequence: 1 aatgaagaccacccctgtgta.....agtccagaaaaagggggg 4141

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1825	44.1	10100	3	US-09-133-944-1
C 2	1825	44.1	10100	3	US-09-208-827-1
C 3	1825	44.1	10100	4	US-10-043-074-1
4	1819.8	43.9	6350	2	US-08-385-335A-8
5	1597.8	38.6	8614	3	US-09-208-827-5
6	1597.8	38.6	8614	4	US-10-043-074-5
7	1497.6	36.2	6620	4	US-08-786-531B-3
8	1496	36.1	5363	4	US-08-786-531B-2
9	1437.8	34.7	8387	2	US-08-532-814-1
10	1437.8	34.7	8388	3	US-09-225-509-1
11	1420.6	34.3	7617	3	US-08-646-538-34
12	1420.6	34.3	7617	3	US-09-503-222-34
13	1374	33.2	8202	1	US-08-258-420-13
14	1230.8	29.7	6700	4	US-09-654-449-1
15	1230.8	29.7	8161	4	US-09-759-152A-1
16	1230.8	29.7	8175	4	US-09-654-449-4
17	1230.8	29.7	8175	4	US-09-654-449-5
18	1230.8	29.7	8518	4	US-09-654-449-3
19	1229.2	29.7	6700	4	US-09-759-152A-1
20	1229.2	29.7	8175	4	US-09-759-152A-6
21	1229.2	29.7	8518	4	US-09-759-152A-4
22	1196.2	28.9	8316	1	US-07-753-520B-4
23	1196.2	28.9	9115	1	US-07-753-520B-3
24	1194	28.8	6365	1	US-08-352-990-1
25	1193.8	28.8	7235	4	US-08-786-531B-6
26	1192.6	28.8	7699	4	US-09-645-004-1
27	1192.6	28.8	7980	4	US-09-645-004-2

28	1192	28.8	5865	3	US-08-654-737B-3
29	1192	28.8	6145	1	US-08-336-132-1
30	1192	28.8	6145	3	US-08-935-312-1
31	1192	28.8	6145	3	US-08-848-760B-1
32	1192	28.8	6145	4	US-09-826-025-1
33	1192	28.8	7086	4	US-09-935-194-1
34	1192	28.8	7160	4	US-08-786-531B-5
35	1192	28.8	7311	4	US-09-645-004-3
36	1192	28.8	7352	4	US-08-786-531B-4
37	1192	28.8	7353	4	US-08-786-531B-1
38	1192	28.8	7885	4	US-09-645-004-4
39	1190.2	28.7	5689	4	US-09-508-516-1
40	1167.8	28.2	10306	3	US-08-716-351A-4
41	1167.8	28.2	10970	3	US-08-716-351A-5
42	1073	25.9	5292	2	US-08-793-610-3
C 43	1051.8	25.4	10367	1	US-08-110-300A-9
C 44	1051.8	25.4	10367	2	US-08-886-642-9
C 45	1051.8	25.4	10367	5	PCT-US93-08041-9

ALIGNMENTS

RESULT 1  
US-09-133-944-1/c  
; Sequence 1, Application US/09133944  
; Patent No. 6280937  
; GENERAL INFORMATION:  
; APPLICANT: Luo, Ying  
; APPLICANT: Yu, Pei Wen  
; APPLICANT: Lorens, James  
; TITLE OF INVENTION: SHUTTLE VECTORS  
; FILE REFERENCE: A66252/DJB/DAV  
; CURRENT APPLICATION NUMBER: US/09/133,944  
; CURRENT FILING DATE: 1999-08-14  
; EARLIER FILING DATE: 1998-08-14  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 10100  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: constructed  
; OTHER INFORMATION: vectors  
US-09-133-944-1

Query Match	44.1%;	Score 1825;	DB 3;	Length 10100;
Best Local Similarity	98.8%;	Pred. No. 0;		
Matches 1859;	Conservative 0;	Mismatches 20;	Indels 2;	Gaps 2;
QY	96	CTAGCTTAAGTAAACCCATTGTCAGGCGATCGAAATACTAGTAAGTATAGAGAG 155		
DB	9382	CTAGCTTAAGTAAACCCATTGTCAGGCGATCGAAATACTAGTAAGTATAGAGAG 9323		
QY	156	TTGAGATCAAGTCAAGGAAACAGATGGAACAGCTGAATATGGCCAAACAGGATATCTGTG 215		
DB	9322	TTGAGATCAAGTCAAGGAAACAGATGGAACAGCTGAATATGGCCAAACAGGATATCTGTG 9263		
QY	216	GTAAACAGTCTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATGGGCCA 275		
DB	9262	GTAAACAGTCTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATGGGCCA 9203		
QY	276	AACAGGATATCTGTGTAAGCAGTTCTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAATGTC 335		
DB	9202	AACAGGATATCTGTGTAAGCAGTTCTCTGCGCCCGGCTCAGGGCCAGAACAGATGGAATGTC 9143		
QY	336	CAGATCGGTCAGCCCTCAGCAGTTTCTAGAGAACCAATCAGATGTTTCCAGGGTGC 395		
DB	9142	CAGATCGGTCAGCCCTCAGCAGTTTCTAGAGAACCAATCAGATGTTTCCAGGGTGC 9083		
QY	396	AAGNACCTGAATACCCCTGTCCTTATTGCACTAACCAATCAGTTGCTGCTTCTCGCTTC 455		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 02:13:00 ; Search time 1326.22 Seconds  
(without alignments)  
18483.824 Million cell updates/sec

Title: US-09-982-223a-2  
Perfect score: 4141  
Sequence: 1 aatgaagacccacactgtg.....agttccagaaaaagggggg 4141

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1980s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4141	100.0	4141	6	ABK85888 DNA seque
2	2143.6	51.8	5715	3	Aaz34936 Retrovira
3	2143.6	51.8	5715	4	Aaf30945 Vector us
4	1840.	44.4	5782	6	ABK85887 DNA seque
5	1834.4	44.3	9037	12	ADG73794 Vector pB
6	1829.4	44.2	8040	4	AAC60482 Vector MF
7	1829.4	44.2	8045	2	Aat12237 Retrovira
8	1829.4	44.2	9629	2	AAT14600 pXJCL-hGM
9	1825	44.1	10100	4	AAS13809 DNA encod
10	1825	44.1	10100	6	ABK15128 Shuttle v
11	1825	44.1	10100	8	ABX95501 Novel shu
12	1823.6	44.0	5903	13	ADQ80674 Bcl-XL ex
13	1823.6	44.0	7257	13	ADQ80673 Bcl-XL ex
14	1819.8	43.9	6350	2	AAT35198 Plasmid p
15	1803.4	43.5	6870	10	ADJ71746 Vector pI
16	1693.8	40.9	7654	8	ABX16565 Retrovira
17	1693.8	40.9	7654	10	ABX13168 Retrovira
18	1597.8	38.6	8614	6	ABK15132 Shuttle v
19	1597.8	38.6	8614	8	ABX95505 Novel shu
20	1585.8	38.3	6221	3	Aaz34935 Retrovira

21	1585.8	38.3	6221	4	AAP30944 Vector us
22	1556.6	37.6	6143	9	ADA12885 Marine re
23	1542.6	37.3	7372	2	AAX33182 Base sequ
24	1541	37.2	6644	2	AAX33181 Base sequ
25	1514	36.6	7797	2	AAX33180 Cowpox vi
26	1509.4	36.5	7996	2	AAX33184 Base sequ
27	1497.6	36.2	6620	6	AD82653 Plasmid v
28	1496	36.1	5363	6	AD82652 Plasmid v
29	1459.6	35.2	8282	3	AAX34938 Cl2ScFas
30	1459.6	35.2	8282	4	AAX34939 Cl2ScFas
31	1459.6	35.2	8345	3	AAX34939 Ahhh sur
32	1459.6	35.2	8345	4	AAX34938 Ahhh sur
33	1458	35.2	4924	3	AAX34937 Retrovira
34	1458	35.2	4924	4	AAX34936 Vector us
35	1445.4	34.9	6253	9	AAL57215 MIG retro
36	1437.8	34.7	8388	2	AAQ78191 Vector M4
37	1429.8	34.5	8889	10	ADH76472 Chimeric
38	1426.6	34.5	11364	10	ADH76479 Chimeric
39	1421.8	34.3	11394	10	ADH76473 Chimeric
40	1420.6	34.3	7617	2	AAV14354 Plasmid p
41	1374	33.2	8202	2	AAT09280 Novel AMP
42	1358.6	32.8	6893	10	ADE24111 Proviral
43	1335.2	32.2	3728	5	AAD04929 Retrovira
44	1335.2	32.2	5365	5	AAD04928 Retrovira
45	1236.8	29.9	6250	4	AAD14296 MESVR/EGF

## ALIGNMENTS

RESULT 1  
ABK85888  
ID ABK85888 standard; DNA; 4141 BP.  
XX  
AC ABK85888;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE DNA sequence encoding Gag packaging protein #2.  
XX  
KW Gag; ds; viral expression vector.  
XX  
OS Unidentified.  
XX  
PN WO200234929-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 18-OCT-2001; 2001WO-US032592.  
XX  
PR 20-OCT-2000; 2000US-0241879P.  
XX  
PA (WBED ) WHITEHEAD INST BIOMEDICAL RES.

PI Kohn EV, Daley GQ;  
XX  
DR WPI, 2002-489949/52.  
XX  
PT Expression vectors comprising viral vectors (e.g. retroviral vectors), useful for generating or screening nucleic acid libraries, or in antisense based or gene trapping methods for identifying modulators of a mammalian gene.  
XX  
PT Claim 6; Fig 2; 114pp; English.  
XX  
PS This invention relates to a novel vector comprising from 5' to 3' a packaging sequence, a heterologous insert sequence or restriction site for insertion of a heterologous sequence and a 3' long terminal repeat (LTR) sequence, where at least 2 codons of the packaging sequence are altered to reduce formation of fusion polypeptides encoded by the packaging sequence or a portion of it, and the heterologous insert sequence. The vector of the invention is useful for generating a library and in screening nucleic acid libraries. In particular, the vector is

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 02:23:00 ; Search time 11478.2 Seconds  
(without alignments)  
17481.226 Million cell updates/sec

Title: US-09-982-223A-2

Perfect score: 4141

Sequence: 1 aatgaagaccacactgtg.....agtctccagaaaaagggggg 4141

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred.. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4141	100.0	4141	6	AX491315	Sequence
2	2131.6	51.5	5713	6	BD222121	Method an
3	2131.6	51.5	5713	6	AX146811	Sequence
4	1840	44.4	5782	6	AX491314	Sequence
5	1834.4	44.3	9037	6	AX951342	Sequence
6	1834.4	44.3	9037	6	AX960754	Sequence
7	1825.2	44.1	5117	12	AF132210	Cloning v
8	1825.2	44.1	6501	12	AF132211	Cloning v
9	1825.2	44.1	11065	12	AF264696	Cloning v
10	1825	44.1	10100	6	AR165918	Sequence
11	1823.6	44.0	5903	6	AR429966	Sequence
12	1823.6	44.0	7257	6	CQ879092	Sequence
13	1823.6	44.0	9609	12	CQ879091	Sequence
14	1822.6	44.0	9609	12	AF396261	Cloning v
15	1819.8	43.9	6350	6	AR069650	Sequence
16	1597.8	38.6	8614	6	AR429970	Sequence
17	1593	38.5	5883	12	AB041927	Retrovira
18	1593	38.5	6248	12	AB041928	Retrovira
19	1585.8	38.3	6219	6	BD222120	Method an

20 1585.8 38.3 6219 6 AX146810  
21 1563.6 37.8 5901 12 AB086388  
22 1563.6 37.8 5904 12 AB086384  
23 1563.6 37.8 6076 12 AB086386  
24 1563.6 37.8 6277 12 AB086385  
25 1563.6 37.8 6824 12 AB086387  
26 1563.6 37.8 6143 6 AX823826  
27 1542.6 37.3 7372 6 E23357  
28 1541 37.2 6644 6 E23356  
29 1516.8 36.6 6690 12 AB086389  
30 1514 36.6 7797 6 E23355  
31 1509.4 36.5 7996 6 E23359  
32 1497.6 36.2 6620 6 AR302094  
33 1496 36.1 5363 6 AR302093  
34 1459.6 35.2 8282 6 BD222123  
35 1459.6 35.2 8282 6 AX146813  
36 1459.6 35.2 8345 6 BD222124  
37 1459.6 35.2 8345 6 AX146814  
38 1446 34.9 4922 6 BD222122  
39 1446 34.9 4922 6 AX146812  
40 1437.8 34.7 8387 6 AR070490  
41 1437.8 34.7 8388 6 AR179512  
42 1429.8 34.5 8989 6 AX768031  
43 1429.8 34.5 11328 12 AF010170  
44 1429.8 34.5 11364 6 AX768038  
45 1429.8 34.5 11394 6 AX768032

## ALIGNMENTS

RESULT 1  
AX491315  
LOCUS AX491315 4141 bp DNA linear PAT 16-AUG-2002  
DEFINITION Sequence 2 from Patent WO0234929.  
ACCESSION AX491315  
VERSION AX491315.1 GI:22324010  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Kohn, E.Y. and Daley, G.Q.  
TITLE Expression vectors and uses thereof  
JOURNAL Patent: WO 0234929-A 2 02-MAY-2002;  
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)  
FEATURES  
source  
1..4141  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Synthetically generated nucleic acid"

## ORIGIN

Query Match 100.0%; Score 4141; DB 6; Length 4141;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AATGAAGACCCACCTCTAGTGTGGCAGCTAGCGCGCGCCATCACTTCGTATAGCA 60  
DB 1 AATGAAGACCCACCTCTAGTGTGGCAGCTAGCGCGCGCCATCACTTCGTATAGCA 60  
QY 61 TACATTATACGAAGTTATTTAATTAAGCGCGCCCTCTAGCTTAAGTAAGCGCATTTTGA 120  
DB 61 TACATTATACGAAGTTATTTAATTAAGCGCGCCCTCTAGCTTAAGTAAGCGCATTTTGA 120  
QY 121 AGGCATGGAATAATACATACTAGAGATAGAGAGTTTCAGATCAAGTTCAGGAACAGATG 180  
DB 121 AGGCATGGAATAATACATACTAGAGATAGAGAGTTTCAGATCAAGTTCAGGAACAGATG 180  
QY 181 GAACAGCTGAATATGGGCGCAACAGGATATCTGTGTAAGCAAGTTCTTCCCGCGCTCAG 240  
DB 181 GAACAGCTGAATATGGGCGCAACAGGATATCTGTGTAAGCAAGTTCTTCCCGCGCTCAG 240

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 08:47:41 ; Search time 11362.4 Seconds  
(without alignments)  
19369.833 Million cell updates/sec

Title: US-09-982-223a-1  
Perfect score: 5782  
Sequence: 1 aatgaagaccacacacgtgta.....tccgcgcacatttcctgcacat 5782

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1014.2	17.5	1070	1	AJ281552
2	976.6	16.9	1048	7	CO552396
3	943.2	16.3	1013	4	BM438846
4	882.4	15.3	1004	1	AJ281480
5	863.8	14.9	902	7	CR753463
6	854.6	14.8	935	4	EG838279
7	848.2	14.7	928	7	CO487414
8	839.4	14.5	841	1	AL042026
9	827	14.3	1073	7	CF269652
10	823	14.2	854	4	BM438950
11	818.8	14.2	865	7	CK125894
12	807.6	14.0	866	8	BZ570738
13	806.8	14.0	856	7	CN823189
14	806.4	13.9	819	6	CD649375
15	803.2	13.9	1126	8	BZ577702
16	793.8	13.7	1249	8	BZ572284
17	791.4	13.7	827	7	CN823902
18	789.4	13.7	1011	8	BZ576726
19	779.8	13.5	846	7	CV468077
20	775.4	13.4	820	7	CR753457
21	773.8	13.4	1067	1	AU081137
22	773.4	13.4	954	1	AL044364
23	773.2	13.4	1336	8	BZ575810
24	772.8	13.4	789	6	CD280920

C 25	769.8	13.3	832	7	CN822433	CN822433	Oa splbn
C 26	764.2	13.2	769	5	CV224987	CV224987	CS hyp 24
C 27	761.6	13.2	780	5	BQ825693	BQ825693	103012380
C 28	758.4	13.1	1574	8	BZ572566	BZ572566	msh2 2693
C 29	755.4	13.1	759	6	CD279661	CD279661	G43818.135
C 30	753.8	13.0	1463	8	BZ571475	BZ571475	msh2 1906
C 31	752	13.0	1370	8	BZ571721	BZ571721	msh2 2025
C 32	749.8	13.0	786	7	CN823164	CN823164	Oa splbn
C 33	747.4	12.9	752	7	CR766850	CR766850	DKFZp468H
C 34	745.4	12.9	844	8	BZ574513	BZ574513	msh2 3706
C 35	743.8	12.9	800	1	AJ281449	AJ281449	4A3A-P4D5
C 36	742.4	12.8	1089	1	AU081124	AU081124	AU081124
C 37	740.6	12.8	863	7	CF752100	CF752100	TGDR9 Hum
C 38	739.2	12.8	741	6	CD279174	CD279174	G44221.83
C 39	739.2	12.8	914	8	BZ569398	BZ569398	pac82-164
C 40	733.2	12.7	811	9	ATH517156	ATH517156	Arabidops
C 41	732.2	12.7	950	8	BZ571129	BZ571129	msh2 1741
C 42	728.4	12.6	730	6	CD281097	CD281097	G44224.38
C 43	728.4	12.6	756	7	CO477685	CO477685	GQ0132.B7
C 44	728.2	12.6	998	8	BZ576702	BZ576702	msh2 5060
C 45	727.6	12.6	793	7	CK122681	CK122681	BES182410

## ALIGNMENTS

RESULT 1	AJ281552	1070 bp	mRNA	linear	EST 30-JUN-2000
LOCUS	4A3A-P6F11-F	Anopheles gambiae	immune competent	4A3A	Anopheles
DEFINITION	gambiae CDNA clone 4A3A-P6F11, mRNA sequence.				
ACCESSION	AJ281552				
VERSION	AJ281552.1	GI:6929432			
KEYWORDS	EST.				
SOURCE	Anopheles gambiae (African malaria mosquito)				
ORGANISM	Anopheles gambiae				
REFERENCE	1 (bases 1 to 1070)				
AUTHORS	Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansong, W., Soares, M.B. and Kafatos, F.C.				
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12)			6619-6624 (2000)
MEDLINE	20300950				
PUBMED	10841561				
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1. 1070 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="4A r/r" /db_xref="taxon:7165" /clone="4A3A-P6F11" /cell_line="immune competent 4A3A" /lab_host="E. coli DH10B" /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."				
FEATURES	source				
ORIGIN					



```

Query Match      100.0%; Score 5782; DB 10; Length 5782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATGAAGACCCCACTGTAGCTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGC AAGC 60
    |||
db 1 AATGAAGACCCCACTGTAGCTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGC AAGC 60

```

Result No.	Score	Match	Query	Length	DB	ID	Description
1	3116.2	53.9	6145	1	US-08-336-132-1	Sequence 1, Appl	
2	3116.2	53.9	6145	3	US-08-935-312-1	Sequence 1, Appl	
3	3116.2	53.9	6145	3	US-08-848-760B-1	Sequence 1, Appl	
4	3116.2	53.9	6145	4	US-09-826-025-1	Sequence 1, Appl	
5	3024.6	52.3	6365	1	US-08-352-990-1	Sequence 1, Appl	
6	2991.4	51.7	6620	4	US-08-786-531B-3	Sequence 3, Appl	
7	2882	49.8	5865	3	US-08-654-737B-3	Sequence 3, Appl	
8	2608.2	45.1	5689	4	US-09-508-516-1	Sequence 1, Appl	
9	2521.2	43.6	7086	4	US-09-935-194-1	Sequence 1, Appl	
10	2490.2	43.1	8518	4	US-09-654-449-3	Sequence 3, Appl	
11	2490.2	43.1	8518	4	US-09-759-152A-4	Sequence 4, Appl	
12	2483.8	43.0	8161	4	US-09-759-152A-8	Sequence 4, Appl	
13	2477.8	42.9	8175	4	US-09-654-449-4	Sequence 4, Appl	
14	2477.8	42.9	8175	4	US-09-654-449-5	Sequence 5, Appl	
15	2477.8	42.9	8175	4	US-09-759-152A-6	Sequence 6, Appl	
16	2454.8	42.5	7160	4	US-08-786-531B-5	Sequence 6, Appl	
17	2411	41.7	7235	4	US-08-786-531B-6	Sequence 6, Appl	
18	2369.6	41.0	5363	4	US-08-786-531B-2	Sequence 2, Appl	
19	2296.2	39.7	7311	4	US-09-645-004-3	Sequence 3, Appl	
20	2263	39.1	7332	4	US-08-786-531B-1	Sequence 1, Appl	
21	2258.8	39.1	7332	4	US-08-786-531B-4	Sequence 4, Appl	
22	2210	38.2	7980	4	US-09-645-004-2	Sequence 2, Appl	
23	2206.4	38.2	7699	4	US-09-645-004-1	Sequence 1, Appl	
24	2206.4	38.2	8316	1	US-07-753-520B-4	Sequence 4, Appl	
25	2206.4	38.2	9115	1	US-07-753-520B-3	Sequence 3, Appl	
26	2205.2	38.1	7885	4	US-09-645-004-4	Sequence 4, Appl	
27	2178.4	37.7	4950	3	US-08-789-333F-58	Sequence 58, Appl	

	Query Match	Best Local Similarity	53.9%;	Score 3116.2;	DB 1;	Length 6145;
	Matches 4510;	Conservative	0;	Mismatches 1078;	Indels 507;	Gaps 24;
QY	3	TGAAGACCCACCTGTAGGTTTGCCAGCTAGCTTAAGTAACGCCATTTTGC	62			
DB	147	TGAAGACCCACCCCGTAGG--TGGCAGCTAGCTTAAGTAACGCCATTTTGC	204			
QY	63	GGRAATATACATCTAGAAATAGAAAGTTTCAGATCAAGGTCAGGAACAGATGAAACAG	122			

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2005, 02:13:00 ; Search time 1851.78 Seconds  
(without alignments)  
18483.824 Million cell updates/sec

Title: US-09-982-223a-1  
Perfect score: 5782  
Sequence: 1 aatgaagaccacactgtg.....tcgcgcacattctctgcac 5782

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5782	100.0	5782	6	ABK85887
2	3657.2	63.3	6221	3	Aax34935 Retrovira
3	3657.2	63.3	6221	4	Aaf30944 Vector us
4	3580.6	61.9	5903	13	ADQ80674 Bcl-XL ex
5	3334.6	57.7	5715	3	Aax34936 Retrovira
6	3334.6	57.7	5715	4	Aaf30945 Vector us
7	3213	55.6	7654	8	ABX16565 Retrovira
8	3213	55.6	7654	10	ABX13168 Retrovira
9	3118.6	53.9	7257	13	ADQ80673 Bcl-XL ex
10	3118	53.9	6444	6	AA167595 Nucleotid
11	3116.2	53.9	6145	2	AAV83182 Intermedi
12	3116.2	53.9	6145	5	AAf85611 M-MuLV-ba
13	3085	53.4	6141	2	AAf309481 Plasmid r
14	3066.2	53.0	6522	2	AAf309482 Plasmid r
15	3023	52.3	6365	2	AAQ41173 Plasmid L
16	2999.8	51.9	6046	12	ADG83262 Clone pLX
17	2991.4	51.7	6620	6	ADG82653 Plasmid v
18	2986	51.6	6644	2	Aax33181 Base sequ
19	2969.6	51.4	6620	9	ADA12887 Murine le
20	2871.6	49.7	5874	2	AAf309484 Plasmid r

21	2862.8	49.5	5865	2	AAV04002 Retrovira
22	2854.6	49.4	5365	5	AAf30946 Vector us
23	2774.6	48.0	6253	9	AAf57215 MIG retro
24	2704.2	46.8	6283	4	AAf83147 Complete
25	2635.4	45.6	6505	9	ADA12886 Murine MS
26	2608.2	45.1	5689	2	AAZ11445 Retrovira
27	2608.2	45.1	5689	3	AAf61061 Retrovira
28	2608.2	45.1	5689	3	AAf96208 Nucleotid
29	2597.8	44.9	7372	2	AAf33182 Base sequ
30	2579	44.6	4924	3	AAZ34937 Retrovira
31	2579	44.6	4924	4	AAf30946 Vector us
32	2521.2	43.6	7086	6	ABK12523 Tetraycycl
33	2490.2	43.1	8518	4	AAh74524 Nucleotid
34	2488.6	43.0	8161	4	AAh74526 Nucleotid
35	2483.8	43.0	8161	4	AAh76190 Nucleotid
36	2477.8	42.9	8175	4	AAh74525 Nucleotid
37	2477.8	42.9	8175	4	AAh76189 Nucleotid
38	2475.8	42.8	7165	2	AAf309483 Plasmid r
39	2454.8	42.5	7160	6	ADG82655 Plasmid v
40	2437	42.1	6143	9	ADA12885 Murine re
41	2419.8	41.9	8518	4	AAh76188 Nucleotid
42	2411	41.7	7235	6	ADG82656 Plasmid v
43	2386.8	41.3	8852	5	AAf55128 Nucleotid
44	2369.6	41.0	5363	6	ADG82652 Plasmid v
45	2352.8	40.7	7840	5	AAf55127 Nucleotid

## ALIGNMENTS

RESULT 1  
ABK85887  
ID ABK85887 standard; DNA; 5782 BP.  
XX  
AC ABK85887;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE DNA sequence encoding Gag packaging protein.  
XX  
KW Gag; ds; viral expression vector.  
XX  
OS Unidentified.  
XX  
PN WO200234929-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 18-OCT-2001; 2001WO-US032592.  
XX  
PR 20-OCT-2000; 2000US-0241879P.  
XX  
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

Kohn EY, Daley GQ;  
WPI; 2002-489949/52.  
Expression vectors comprising viral vectors (e.g. retroviral vectors),  
useful for generating or screening nucleic acid libraries, or in  
antisense based or gene trapping methods for identifying modulators of a  
mammalian gene.

Claim 8; Fig 1; 114pp; English.

This invention relates to a novel vector comprising from 5' to 3' a  
packaging sequence, a heterologous insert sequence or restriction site  
for insertion of a heterologous sequence and a 3' long terminal repeat  
(LTR) sequence, where at least 2 codons of the packaging sequence are  
altered to reduce formation of fusion polypeptides encoded by the  
packaging sequence or a portion of it, and the heterologous insert  
sequence. The vector of the invention is useful for generating a library  
and in screening nucleic acid libraries. In particular, the vector is

Result No.	Query Match	Score	Length	DB	ID	Description
1	100.0	5782	6	AX491314	Sequence	
2	3948	5883	12	AB041927	Retrovirus	
3	3922.4	6076	12	AB086386	Retrovirus	
4	3897.2	67.8	6248	AB041928	Retrovirus	
5	3802.8	65.8	5904	AB086384	Retrovirus	
6	3802	65.8	5901	AB086388	Retrovirus	
7	3794.2	65.6	6277	AB086385	Retrovirus	
8	3786.2	65.5	6501	AF133211	Cloning	
9	3633.4	62.8	6219	BD222120	Method	
10	3633.4	62.8	6219	AX146810	Sequence	
11	3580.6	61.9	5903	CQ879092	Sequence	
12	3389.2	58.6	6690	AB086389	Retrovirus	
13	3310.6	57.3	5713	6 BD222121	Method	
14	3310.6	57.3	5713	AX146811	Sequence	
15	3279.2	56.7	6824	12 AB086387	Retrovirus	
16	3174.4	54.9	6374	12 SYNMOV1	Mooney	
17	3163.8	54.7	6259	12 SYNRRV	Cloning	
18	3119.4	54.0	6149	12 SYNMMPLN2	Cloning	
19	3118.6	53.9	7257	6 CQ879091	Retroviral	

## ORIGIN

Query Match	100.0%	Score 5782;	DB 6;	Length 5782;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5782;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1 AATGAAAGACCCACCTGTAGCTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGCAGGC	60			
1 AATGAAAGACCCACCTGTAGCTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGCAGGC	60			
61 ATGGAAAATAACATACTGAGAATAAGAAAAGTTTCAGATCAAGGTTCAGGAACAGATGGAAC	120			
61 ATGGAAAATAACATACTGAGAATAAGAAAAGTTTCAGATCAAGGTTCAGGAACAGATGGAAC	120			
121 AGCTGAATATGGGCCAAGCGGATATCTGTGTAGACGATTCCTGCCCGGCTCAGGGCC	180			
121 AGCTGAATATGGGCCAAGCGGATATCTGTGTAGACGATTCCTGCCCGGCTCAGGGCC	180			
181 AAGAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGACGATTCCTCG	240			
181 AAGAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGACGATTCCTCG	240			